



Universidad Nacional de La Plata
Facultad de Bellas Artes
Taller de Diseño Multimedial 5
Tesina de Grado

Diseño de Interfaces Gestuales, aplicado a la Instalación Multimedia

Profesor: Federico Joselevich Puiggrós
Alumna: Jérica Toledo, legajo nº 57502/4

-2014-

ÍNDICE

RESUMEN	3
INTRODUCCIÓN	3
GESTOS	5
COMUNICACIÓN NO VERBAL	6
INTERACCIÓN DIRECTA E INDIRECTA	7
DISEÑO DE INTERFACES GESTUALES	8
CONCEPTO	10
CONCLUSIÓN	12
BIBLIOGRAFÍA	13

Diseño de Interfaces Gestuales, aplicado a la Instalación Multimedia

“La nueva era del diseño de interacción
se mueve por la ausencia de paradigmas
que nos imponían el teclado y el mouse”.

Dan Saffer ¹

Resumen

Si las expresiones dan información de quien las acciona a quien las recibe, la comunicación no verbal tiene como fuerte las expresiones gestuales de cada individuo. La comunicación computarizada a través de gestos la denomino con el nombre de comunicación no verbal computador-humano. Clave importante en este análisis, que resume mi tesis como interfaces gestuales que captan los movimientos físicos de las personas en un espacio y en un contexto determinado.

Introducción

Actualmente nos encontramos con un gran avance en la Inteligencia Artificial (IA). Uno de los campos que incluye en su investigación, desarrollo y producción es la relación que se establece en la comunicación humano-máquina. En este punto me voy a centrar, sumado a un breve análisis del estudio de las interfaces gestuales. Las interfaces gestuales se han vuelto el centro de atención de muchas empresas, artistas y consumidores de nuevas tecnologías. Estas interfaces tienen como principal objetivo detectar los gestos faciales y corporales del ser humano. Como ejemplo de este argumento podemos ver cómo la publicidad BTL de Samsung Mobile con el Galaxy 4S, hace el lanzamiento de este producto con la tecnología de Air View, permite mover el dedo cerca de la pantalla sin tocarla y que el dispositivo ejecute acciones. A su vez, incluye "Smart Scroll que para navegar por una página en el dispositivo a través del reconocimiento ocular de cada usuario, con lo que se podrá controlar el aparato con el movimiento de los ojos", InformaBTL Promociones, activaciones y below the line (<http://www.informabtl.com/>).



Creativa acción BTL que no debes perder de vista (All eyes on the Samsung S4)

En esta publicidad las personas se aproximaban a un monolito con la forma del celular Samsung S4 con la tecnología Smart Scroll, ofreciendo de regalo su producto a aquella persona que logre mirar continuamente un minuto la cámara incorporada al monolito.

Para desarrollar el diseño de interfaces gestuales es muy importante saber qué se quiere realizar, para qué va a ser destinada esa interfaz y cómo se introducirá al contexto elegido previamente. A la hora de introducir dicha tecnología al campo artístico, es recomendable elaborar un diseño específico de la interfaz gestual y con ello, ver qué tecnologías ya existentes colaboran con los requisitos de la obra a realizar.

El diseño de interfaces para las interacciones gestuales, comúnmente se lleva a cabo por gestos que en cierta forma son convencionales, pensados desde la interfaz gráfica hasta la interfaz física, estipulando el contexto, determinando la cantidad de personas que podrán hacer uso simultáneo de la misma, flujo de feedback (acción-reacción) entre la acción gestual y la respuesta.

Considero que hemos llegado al punto donde la tecnología está mostrando que no es necesario tocar físicamente un objeto para jugar, investigar, diseñar, etc. Las obras interactivas que incorporen interfaces gestuales buscando traspasar el tacto y controlar objetos físicos realizando gestos en el espacio, lograrán un lenguaje interactivo intuitivo como signos de la obra situada en la contemporaneidad, como afirma Claudia Giannetti en *El Espectador Como Interactor*, "La interactividad pasa a ser signo no sólo de contemporaneidad de la obra, sino incluso de su cualidad".

El arte generativo cuenta con el desarrollo de tecnología que se adapta al concepto de una obra artística y en particular diseña, programa, genera todo lo necesario para poder dar vida a un objeto, una instalación, una obra en sí misma conceptual creada a partir de la tecnología existente o descubre nuevas tecnologías con el desarrollo de ese trabajo.

En estos últimos años la robótica junto con la Computación Afectiva que fundó Rosalind Picard, Ingeniera Electrónica, se ha interesado en la investigación de las experiencias de las personas con

dispositivos desde las emociones, en sus palabras "la informática que se relaciona con las emociones , no sólo con las consideradas más importantes, como la alegría o la tristeza, sino también con el interés, el aburrimiento o la frustración, que son las que se dan en relación con los ordenadores".²

GESTOS

El estudio de los gestos se centra en la comunicación no verbal que las personas tienden a realizar mediante acciones corporales. La serie televisiva estadounidense "Lie to me" (Miénteme), se basa en la investigación de Paul Ekman, psicólogo pionero en el estudio de las emociones y la expresión facial de las mismas. Se presenta en concepción a la investigación de los siete gestos universales de las personas, ellos son: alegría, ira, miedo, repugnancia, sorpresa y tristeza; luego su lista de gestos básicos se amplió con su continuo estudio de las emociones y los gestos faciales.³



LIE TO ME

Las expresiones se incorporan de persona a persona, en un mismo núcleo social y cultural, las personas interactúan y expresan sus emociones, no necesariamente a través del diálogo, muchas veces el rostro representa una respuesta o una acción que comunica a su receptor.

Los gestos de las personas son el fuerte de su comunicación, es decir que afirman, asienten, niegan, refuerzan lo que dicen con sus palabras. En la vida cotidiana los usamos sin pensar, sólo actuamos según se van dando las circunstancias.

En esta línea de investigación también se analiza la incorporación de las manos, ya que con ellas construimos, hacemos las cosas, sin ellas es imposible manejarnos, y gracias a ellas es posible comunicarnos de manera no verbal.

En una instalación multimedia podemos combinar una serie de tecnología para hacer posibles muchísimas obras con la interacción gestual. Hagamos un repaso, la importancia del querer hacer y el hacer posible esa interacción gestual en tiempo real requiere un diseño previo. Establecer los signos, lenguajes, disposiciones del interactor y de los dispositivos, etc. En este marco de elección urge el compromiso de diseñar esas interfaces que ya existen en nuestro entorno y aplicarlas al proyecto a realizar.

COMUNICACIÓN NO VERBAL

En este apartado encontramos una relación entre la comunicación entre personas y la que se puede producir entre computador-humano. Es similar, pero no igual. Con esto hago referencia a que los modelos mentales, por ejemplo una persona saluda a la distancia a otra persona moviendo una de sus manos o una persona usa el mismo saludo hacia una web-cam. Acciones preestablecidas de la vida cotidiana se utilizan en esta relación de interacción computador-humano.

Este tipo de comunicación es a través de signos y símbolos expresados por el cuerpo humano y se refuerza con el contexto que lo rodea, es una comunicación muy común entre las personas. Pero su estudio es muy reciente y carece de una base teórica sólida, sólo voy a nombrar sus usos y sistemas comunes de empleo.

El sistema de comunicación no verbal según Ana María Cestero Mancera, profesora titular de la Universidad de Alcalá-España ⁴, se puede clasificar por tres signos no verbales según sus usos:

- Usos sociales: para la interacción social.
- Usos estructuradores de discursos: para organizar la comunicación e interacción.
- Usos comunicativos: para expresar sensaciones, sentimientos y determinar nociones.

La autora en su artículo “La comunicación no verbal y el estudio de su incidencia en fenómenos discursivos como la ironía”, enumera cuatro sistemas de comunicación no verbal reconocidos hasta el momento:

-El para-lenguaje, se relaciona con las cualidades físicas del sonido y tipo de voz, aportando cualquier expresión corporal y sugieren interpretaciones propias de la lingüística.

-La quinésica, es conocida por la comunicación corporal acompañada de gestos no orales, de percepción visual, auditiva o táctil, por medio de gestos emblemáticos, ilustrativos, reguladores de la interacción, estados emotivos y de adaptación.

Estos dos primeros sistemas son los más comunes y los denomina como “básicos o primarios” porque funcionan a la par cuando una persona está hablando. Luego, la autora presenta a los sistemas secundarios o culturales:

-La proxémica, relacionada con los hábitos, el contexto cultural y creencias de las personas, en base al uso del espacio y uso de la distancia en relación con las personas, entendida como proxémica interaccional.

-La cronémica, se vincula a 'la concepción, estructuración y uso del tiempo que hace el ser humano'.⁵, es decir que el tiempo comunica pasivamente, en la información cultural o activamente que modifica o refuerza el significado de la comunicación. Se puede dar tres tipos de tiempo: conceptual, social e interactivo.

INTERACCIÓN DIRECTA E INDIRECTA

Hay tecnología que capta los gestos a través del tacto, en la tecnología táctil, nosotros manipulamos los objetos digitales tocando pantallas que son sensibles al tacto y también podemos manipular objetos físicos de la realidad sin intervenir por medio del tacto, es decir que podemos encontrarnos con dos tipos de interacción gestual:

-Directa. Es la manipulación a través del tacto sobre una superficie, directamente se interviene un objeto. Quizás ya es común para muchas personas interactuar de esta manera, se aceptaron los códigos del sistema de comunicación entre usuario y dispositivo, ya hay modelos mentales que son fáciles de incorporar en la experiencia de las personas, me refiero al uso de pantallas táctiles para tabletas, celulares, cajeros, GPS, LCD, etc.

-Indirecta. Se diferencia de la directa porque la manipulación gestual se realiza en el espacio y es captada a través de sensores. No necesita una superficie.

Una instalación que me parece interesante en relación al Arte Robótico es “Vincent & Emily”⁶, creada por Nikolas Schmid-Pfähle, se trata de una pareja que interactúa cuando se encuentra sola, pero interactúa de otra manera cuando intervienen personas en su contexto. Esta pareja se representa en un cuarto con dos robots que se atraen entre sí pero se alejan cuando identifican el acercamiento de personas y cada uno de los robots empatiza con ellas de manera individual detectando los gestos y acciones que realicen al momento de acercarse. En este ejemplo podemos ver como las personas interactúan sin tocar los robots, el medio de interacción es la presencia.



VICENT & EMILY

Pranav Mistry, es el inventor de Sixthsense, crea una interfaz gestual para aumentar nuestro mundo. En este caso, se crean una especie de cuatro “dedales” para los dos índices y los dos pulgares respectivos de la mano izquierda y derecha. Su interacción está en una línea muy fina entre la interacción directa en indirecta, porque la idea de este proyecto “es convertir cualquier superficie en una pantalla táctil”⁷, pero ello no descarta que la interacción será indirecta, generando movimientos en el aire sin la necesidad de apretar un botón físico para poder interactuar con la interfaz gráfica.



PRANAV MISTRY

DISEÑO DE INTERFACES GESTUALES

En este apartado, mi interés es interpretar por qué es importante hacer un diseño de las interfaces físicas y gráficas de un dispositivo de interacción gestual.

-Desde na mirada técnica:

La detección de los gestos corporales y faciales se pueden agrupar en dos:

.cara: los gestos faciales tienen muchas zonas que expresan las emociones más comunes: ojos (párpados y pupila), boca (labios y dientes), nariz, cejas, mejillas, frente y mentón.

.cuerpo: se detectan las zonas del cuerpo por las articulaciones, con ellas se puede identificar la cabeza, el cuello, el tórax, los brazos, las manos (y los dedos), las piernas y los pies.

Hay muchos programas (software) y dispositivos (hardware) que fueron creados con el propósito de detectar estímulos corporales y faciales. En los últimos años, gracias a la difusión masiva de información por internet, pudimos conocer proyectos de dispositivos basados en la detección, muy intuitivos, de mayor seguimiento y apoyo a su producción mediante la financiación de grandes empresas interesadas a patrocinar estos emprendimientos.

.Programas: A medida que se fueron acercando a los patrones de detección del cuerpo, en paralelo se fue refinando la detección de las zonas más expresivas y sensibles del rostro humano, generando softwares que mediante un sistema de 'tracking' se pueden establecer los gestos mas básicos.

Un sistema de tracking o detección de personas a través de visión por computador es como dice Yáñez García, Javier en Tracking de personas a partir de visión artificial: "(...)consigue la detección de personas en cada frame mediante un algoritmo que realiza la distinción entre primer plano en movimiento (del inglés, foreground) y el fondo (del inglés, background). De esta manera se consigue aislar los píxeles en los que hay movimiento del fondo, consiguiendo detectar los objetos que se mueven." (fuente: <http://e-archivo.uc3m.es/handle/10016/10118>).

Hay un estudio muy interesante sobre el sistema de detección facial mediante una WebCam y una base de imágenes que aportan a los reconocimientos de cada zona del rostro. Este estudio fue realizado por Cao Chen, Qiming Hou, Kun Zhou, con el título "Displaced Dynamic Expression Regression for Real-time Facial Tracking and Animation", es de fácil acceso y contribuye en gran manera al proceso de detección, como dice su título, es un tracking de rostro en tiempo real.

.Dispositivos: Como ejemplo de este interés mundial, para que la interacción mediante gestos sea más amigable, vemos como "Smart Ring Fin" anillos con Bluetooth, fueron creados con el fin de controlar nuestros dispositivos, desde un Smartphone hasta la televisión inteligente o los sistemas de audio de nuestro automóvil. Todo ello mediante conexión Bluetooth y sencillos gestos. Dichos gestos están detallados en la web del producto.⁸

Los nuevos y populares Smart Tv, las nuevas pantallas inteligentes se volvieron como atracción del momento, que funcionan por comando de voz o gestos, queremos controlar todo sin la necesidad de tener un objeto que medie para que haga algo, entre ellos se encuentran LG y Samsung, con una trayectoria interesante en mejorar cada día con sus productos.⁹

Myo, brazalete inteligente que detecta el comportamiento de los estímulos que los músculos del brazo hacen cuando realizan movimientos. Esta información se almacena en una base de datos y con ellos se puede re-programar un código para generar la dirección del puntero de la computadora, cambiar de canal, manipular objetos como un reproductor de música, etc. Myo es un producto en desarrollo, a cargo de un grupo de investigación llamado Thalmic Labs, y uno de sus proyectos más interesantes es el control de un dron.¹⁰

Con respecto a la captura corporal podemos encontrarnos con dos dispositivos que han revolucionado el mercado internacional de la tecnología: El primero es Kinect un controlador que fue creado para la video consola Xbox 360 y luego para PC bajo el sistema operativo de Windows 7 y 8, fabricado por Microsoft. "El dispositivo cuenta con una cámara RGB, un sensor de profundidad, un micrófono de múltiples matrices y un procesador personalizado que ejecuta el software patentado, que proporciona captura de movimiento de todo el cuerpo en 3D, reconocimiento facial y capacidades de reconocimiento de voz."¹¹ El objetivo de este controlador es capturar y sensar uno o varios cuerpos, comando de voz, creado para jugar sin tener que tocar otros dispositivos¹². El segundo dispositivo es el Leap-Motion, es un sensor para la detección, sensado y captura de los movimientos físicos que realizan las manos en el espacio.¹³

-Desde una mirada artística:

Los estudios sobre la capacidad de dispositivos para captar estímulos gestuales han sido incluidos al arte con modelos mentales, metáforas para la construcción de interfaces gestuales y métodos de aprendizaje. Desde hace más de 30 años vemos como fue evolucionando el arte media incorporando IA, robótica, y a su vez el sentimiento y la emoción conocida como el campo de la Human Computer Interaction (HCI) y la Computación Afectiva (AC, Affective Computing). La AC “es una disciplina de la Inteligencia Artificial que intenta desarrollar métodos computacionales orientados a reconocer emociones humanas y generar emociones sintéticas”.¹⁴

Se puede intervenir sin tocar, en una obra artística esto es difícil de lograr por el campo de acción de la interfaz, la conectividad y la interacción, pero como vemos a continuación no fue imposible en “Very Nervous System” de 1986-90, el artista David Rokeby¹⁵ creó un espacio sonoro en el que una persona danza y sus movimientos corporales emiten sonidos componiendo una pieza musical. En esta forma de interactuar, vemos que no es necesario intervenir un objeto de forma táctil para que accione. Por otro lado Imogen Heap una reconocida artista británica, crea música electrónica sintetizada, es cantante y hace sus shows en base a la composición musical en vivo. Esto se hace posible gracias a Mi-Mu Gloves¹⁶ unos guantes electrónicos, que fueron inventados para este motivo, y son capaces de crear y sintetizar los sonidos. La artista programa y usa una base de datos con sonidos predefinidos por ella, moviendo su manos en el espacio. Más allá que su medio para crear esta interacción sea tener unos guantes, las acciones ocurren en el espacio, los gestos son producidos en el aire.

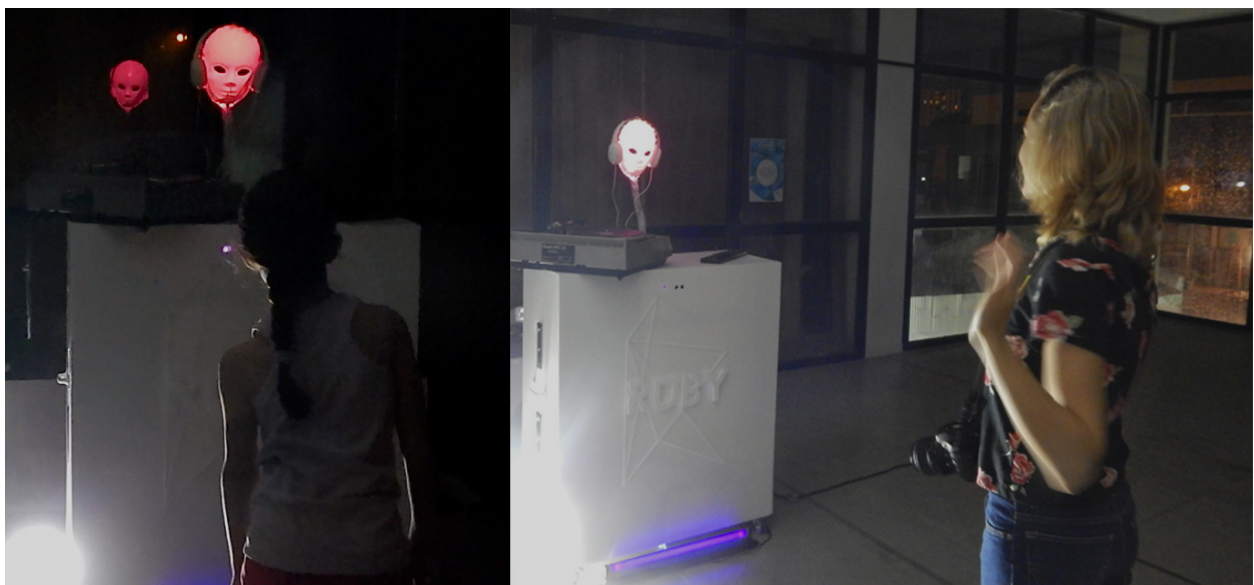
CONCEPTO

En el proceso de una instalación multimedia que lleva como objeto artístico las interfaces gestuales con interacción indirecta, es ineludible pensar por qué dicha tecnología es mejor que aquella que se realiza por tacto. Es todo un desafío plantear una instalación con estas características. El lenguaje conceptual está muy vinculado a las obras contemporáneas que muestran que la interacción tiene que establecer los comportamientos del interactor. Es decir, que el artista creador de dicha instalación deberá conocer al público al que se dirige (ya sea: grupo social, cultural, político). Todo ello influye a la hora de construir un lenguaje artístico que pueda funcionar. En mi instalación multimedia Ruby, investigué sobre estos ejes, para lograr que en ella todas las personas que asistan al lugar de la instalación la puedan entender e inclusive participar.



RUBY

Ruby es un VJ (Video Jockey) interactivo, el cual es capaz de capturar los gestos físicos de una persona y generar mezclas audiovisuales en tiempo real. Las emociones y acciones físicas se ven representadas bajo un fuerte sistema de comunicación no verbal. Ruby espera que lo mires, que logres una conexión con él. Fue creado para influenciar emocionalmente a las personas, también tiene emociones y las expresa. Su expresión se ve reflejada por una mezcla multimedial, arte actual que se caracteriza por la combinación de imagen y sonido.



Su capacidad de detección es inmediata, toma como referencia la detección de la persona que se posiciona frente a él, capturando primeramente el rostro, luego manos y cuerpo. A partir de ahí procesa la información del movimiento de la persona y la traduce a un V Jing a tiempo real (mezcla audiovisual). Si la persona se mueve en el eje 'x', la visual se modifica en un tracking y si se mueve en el eje 'y', la detección se actualiza y queda en loop el último tramo audiovisual hasta que la persona sea otra vez reconocida.

Lo que me interesa resaltar es que es posible diseñar una instalación que aplique la tecnología de detección gestual e implementar un lenguaje fácil de comprender para el espectador, desde los nuevos modelos mentales que hay en nuestra actualidad. Su disposición conceptual y de armado técnico se vinculan, reforzando aún más la interacción. Como observación a esto último quiero agregar que participaron personas de todas las edades. Se involucraron con Ruby produciendo así un espacio audiovisual relacionado con la cultura del V Jing.

CONCLUSIÓN

Todas las emociones, sensaciones y acciones están condicionadas por nuestro contexto virtual en redes sociales, y el arte multimedia incorpora estas pautas para acompañar estos múltiples significados que se le aportan a los dispositivos, como dice Andrés Rodrigues “(...)los artefactos tienen múltiples niveles de significado”¹⁷ y se diseñan a partir de múltiples necesidades y para usuarios específicos. Por eso, es importante que en el Diseño de Interfaces Gestuales aplicados a la Instalación Multimedia, conocer al receptor e identificar su contexto de uso.

Concluyendo así con mi tesis, puedo decir que los gestos de las personas van a ser ejecutados siempre de formas similares, por eso la tecnología puede captar estos gestos y diseñar interfaces que se amolden a todo tipo de movimientos corporales y faciales. Las interfaces gestuales trascienden la interacción por tacto, controlando objetos físicos realizando gestos en el espacio logrando una mayor fluidez interactiva. Siguiendo la concepción de Dan Saffer estamos en una era de propagación interactiva y comunicación constante en el mundo virtual.

~

BIBLIOGRAFÍA

- 1 Nelson Rodríguez-Peña, “Gestuales, tangibles y de cuerpo entero: nuevas interacciones”. Art: “Entrevista a Dan Saffer” pág 8-15, pub. N4 de Faz, Revista de diseño de interacción.
http://www.revistafaz.org/n4/entrevista_dan_saffer.pdf
- 2 R.W. Picard, Esta entrevista fue publicada en julio de 2003, en el número 266 de MUY Interesante.
<http://www.muyinteresante.es/historico/articulo/rosalind-picard>
- 3 Paul Ekman, Professor Emeritus in Psychology at UCSF, Film “Lie to me”.
<http://www.paulekman.com/lie-to-me/>
- 4 Ana María Cestero Mancera, “La comunicación no verbal y el estudio de su incidencia en fenómenos discursivos como la ironía”, Universidad de Alcalá.
http://rua.ua.es/dspace/bitstream/10045/6074/1/ELUA_20_03.pdf
- 5 Definición: “Cronémica”, Enciclopedia Wikipedia.
<http://es.wikipedia.org/wiki/Cron%C3%A9mica>
- 6 Nikolas Schmid-Pfähle, “Vincent & Emily”, Instalación Interactiva 2012.
<http://www.nachtproduktion.de/vincentundemily>
- 7 Pranav Mistry, inventor de Sixthsense.
<http://www.pranavmistry.com/projects/sixthsense/>
- 8 “Smart Ring Fin”.
<http://www.wearfin.com/>
- 9 Smart Tv. LG y Samsung.
<http://www.lg.com/es/smart-tv>
<http://www.samsung.com/es/consumer/av/smart-signage-tv/smart-signage-tv/>
- 10 Thalmic Labs, Myo.
<https://www.thalmic.com/about/>
- 11 Descripción técnica: Kinect, Enciclopedia Wikipedia.
<http://es.wikipedia.org/wiki/Kinect>
- 12 Kinect for Xbox 360. Especificaciones.
<http://www.xbox.com/en-us/kinect>
- 13 Leap Motion. Especificaciones.
<https://www.leapmotion.com/>
- 14 E. Causa y A. Sosa, “La computación afectiva y el arte interactivo” pág 52-60, RIM - Nro. 2, Revista de Investigación Multimedia, Artes Multimediales.
<http://multimedia.iuna.edu.ar/userfiles/publicaciones/archivos/rim2.pdf>
- 15 David Rokeby, “Very Nervous System” of 1986-90.
<http://www.davidrokeby.com/vns.html>
- 16 Imogen Heap, artista británica y una de las fundadoras de Mi.Mu Gloves.
<http://www.imogenheap.co.uk/>
<http://mimu.org.uk/>
- 17 Andrés Rodríguez, “El gesto como mecanismo de interacción corporizada con computadoras: posibilidades y desafíos”, Actas de la IX Reunión.
http://www.saccom.org.ar/2010_reunion9/actas_indice.html

---Otras referencias:

- R.W. Picard (1997), "Affective Computing," MIT Press, Cambridge, 1997.

<http://mitpress.mit.edu/books/affective-computing>

- Cao Chen, Qiming Hou, Kun Zhou: "Displaced Dynamic Expression Regression for Real-time Facial Tracking and Animation", ACM Transactions on Graphics (SIGGRAPH), 2014.

<http://gaps-zju.org/DDE/>